A SYSTEM-THEORETIC VIEW OF BEHAVIOR MODIFICATION

by

L. A. Zadeh

Memorandum No. ERL-M320

January 18, 1972

A SYSTEM-THEORETIC VIEW OF BEHAVIOR MODIFICATION

by

L. A. Zadeh

Memorandum No. ERL-M320

January 18, 1972

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A System-Theoretic View of Behavior Modification*

L. A. Zadeh**

To someone like myself, steeped in the quantitative analyses of inanimate systems, the principal ideas in Skinner's Beyond Freedom and Dignity [1] are difficult to translate into assertions which are capable of proof or refutation. Nevertheless, I find them highly interesting and thought-provoking.

It is a truism that human behavior is vastly more complex than the behavior of man-conceived systems. Reflecting this fact, such basic concepts as control, reinforcement, feedback, goal, constraint, decision, strategy, adaptation, environment, etc., which are central to the discussion of human behavior, are much better understood and more clearly defined in system theory - which deals with abstract systems from an axiomatic point of view - than in psychology or philosophy. Unfortunately, high precision is rarely compatible with high complexity. Thus, the precision and determinism of system theory have the effect of severely restricting its capability to deal with the complexities of human behavior.

Essentially, inanimate systems are amenable to quantitative analysis because their behavior is sufficiently simple to admit of characterization by equations involving numerical variables, i.e., scalars or vectors whose

*Presented at the Conference on "The Social and Philosophical Implications of Behavior Modification," The Center for the Study of Democratic Institutions/The Fund for the Republic, Inc., Santa Barbara, Calif., January 17-19, 1972.

**Department of Electrical Engineering and Computer Sciences, and the Electronics Research Laboratory, University of California, Berkeley, California 94720. This work was supported in part by the NSF Grant GK-10656X.

components are real or complex numbers. Typically, the state of an inanimate system S at time* t, $t = 0,1,2,\ldots$, is an n-vector, x_t, of low or moderate dimensionality, whose components are real numbers. For example, if S is a point of mass m moving in a three-dimensional space, then its state has six components of which the first three define its position and the last three- its velocity.

If S is subjected to a sequence of inputs, u_o, u_1, $u_2,\ldots$, each of which is a numerical variable, then the behavior of S is usually characterized by two equations

$$x_{t+1} = f(x_t,u_t) \tag{1}$$

$$y_t = g(x_t,u_t) \tag{2}$$

The first equation defines the next state (i.e., the state at time t + 1) as a function of the present state, x_t, and the present input, u_t. The second defines the present output, y_t, as a function of the present state and the present input. Thus, the behavior of a deterministic discrete-time system may be characterized by two functions f and g which define, respectively, the next state and the output of the system.

In the past, attempts to describe human behavior by equations of the form (1) and (2) have met with little success because human behavior, in general, is much too complex to admit of description by numerical variables. However, as suggested in [2] and [3], a possible way of dealing with the problem of complexity is to employ <u>fuzzy</u> variables - in place of numerical variables - in (1) and (2). In the case of such variables, their values

---

*For simplicity, we assume that time varies discretely. Dependence on t will frequently be assumed but not exhibited explicitly.

are not numbers but labels of fuzzy sets,* that is, names of classes which do not have sharply defined boundaries. For example, the terms green, big, tired, happy, young, bald, oval may be viewed as labels for classes in which the transition from membership to non-membership is gradual rather than abrupt. Thus, a man aged 32 may have partial membership - represented by a number, say, 0.6 - in the class of young men. The class of young men, then, would be characterized by a membership function $\mu_{young}(x)$ which associates with each man x his grade of membership in class of young men. For simplicity, membership functions are assumed to take values in the interval [0,1], with 0 and 1 representing non-membership and full membership, respectively.

The use of fuzzy variables to describe human behavior is, in effect, a retreat into imprecision in the face of complexity. This, of course, is what has been done all along in psychology and philosophy. However, the use of fuzzy variables in conjunction with equations such as (1) and (2) may make it possible to deal with human behavior in a more systematic and somewhat more precise fashion than is customary in psychology and related fields.

In what follows, we shall sketch the rudiments of this approach and relate it, in part, to human behavior modification. In our brief discussion of the equations characterizing human behavior, we shall not attempt to specify the functions of fuzzy variables which enter into these equations, nor shall we concretize the meaning of the variables representing state, input, environment, etc. Thus, our very limited aim in the present paper is

*Intuitively, a fuzzy set is a class with unsharp boundaries. More precisely, a fuzzy set A in a space $X = \{x\}$, is a collection of ordered pairs $A = \{(x, \mu_A(x))\}$, in which $\mu_A(x)$ is the grade of membership of x in A, with $0 \leq \mu_A(x) \leq 1$. A more detailed discussion of fuzzy sets may be found in [2] and [3].

merely to suggest that some of the aspects of behavior modification which are discussed by Skinner may be formulated, perhaps more systematically, through the use of equations and functions involving fuzzy, rather than numerically-valued, variables. It should be understood, of course, that the detailed task of characterizing the functions entering into these equations by tables or flow-charts involving labels of fuzzy sets would normally require a great deal of psychological testing and data analysis.

Our point of departure is the assumption that the behavior of a human - who for convenience will be referred to as H - can be represented, in part, by the following two pairs of equations

(3) $$x_{t+1} = h_1(x_t, u_t, e_t, t)$$

(4) $$y_t = h_2(x_t, u_t, e_t, t)$$

(5) $$s_{t+1} = g_1(s_t, u_t, y_t, t)$$

(6) $$e_t = g_2(s_t, u_t, y_t, t)$$

in which

$x_t \triangleq$ state of H at time t, $t = 0,1,2,\ldots$

$u_t \triangleq$ action taken by H at time t, with u_t chosen from a constrained (possibly fuzzy) set of alternatives

$e_t \triangleq$ input representing the effect of the external influences not under the control of H (e.g., the effect of the environment, both physical and social)

$y_t \triangleq$ response of H to action u_t and external influences e_t

$s_t \triangleq$ state of environment at time t

$h_1, h_2, g_1, g_2 \triangleq$ fuzzy and, possibly, random functions.

It is understood that some or all of the variables in the above equations are fuzzy, which means that their values are labels of fuzzy sets, e.g., x_t = tired, u_t = taking a nap, e_t = hot and humid, etc. Thus, a typical entry in a table characterizing (3), say, would read, in words: If at time t the state of H is a fuzzy set described by a label α (e.g., α = tired); the effect of the environment is a fuzzy set described by a label β; and the action taken by H is a fuzzy set labeled γ; then with high likelihood the next state of H will be a fuzzy set labeled δ; and possibly, but much less likely, the next state will be ε.

In effect, the first pair of equations, (3) and (4), serves to describe in a very approximate, and yet systematic, fashion the response of H (or some particular aspect of the response of H, represented by y_t) to the external influences (represented by e_t) and the action taken by H (represented by u_t). In a similar fashion, the second pair describes the effect of the behavior of H on the environment. Generally, the effect of H on the environment is much smaller than the effect of the environment on H. This is not true, however, in the case of operant conditioning, where the changes in environment serve to reinforce a particular mode of behavior of H.

To make the description of the behavior of H more explicit, we need an additional equation which describes the decision principle employed by H in selecting an action u_t from a constrained set of alternatives. To this end, it is expedient to make use of the notion of the <u>maximizing set</u> of a function, which is an approximation to - or, in our terminology, a fuzzification of - the notion of a maximizing value.

Suppose that f(x) is a real-valued function which is bounded both from below and from above, with x ranging over a domain X. The maximizing set of f is a fuzzy set, M, in X such that the grade of membership, $\mu_M(x)$, of x in M represents the degree to which f(x) is close to the maximum value of f over X, that is, Sup f (Sup f = supremum of f(x) over X.) For example, if at $x = x_1$ $\mu_M(x_1) = 0.8$, then at $x = x_1$ the value of f(x) is about 80% of its maximum value with respect to some reference point. In effect, then, the maximizing set of a function, f, serves to grade the points in the domain of f according to the degree to which f(x) approximates to Sup f.*

Now let $R_t(u_t)$ denote the estimated total reward** associated with action u_t at time t, with the negative values of R_t representing loss, pain, discomfort, etc. Then we postulate that the decision principle employed by H is the following: For each t at which a decision has to be made, H chooses that u_t which is the maximizing set for the estimated reward. It is understood that, if the membership function of this set does not peak fairly sharply around some particular action, then H first narrows his choice to those actions which have a high grade of membership in u_t and

---

*In more precise terms, the membership function of the maximizing set of a real-valued function f(x), $x \in X$, is defined by the following equations. (Inf f = Infimum of f(x) over X.)

$$\mu_M(x) = \frac{f(x)}{\text{Sup } f} \quad \text{if} \quad \text{Inf } f \geq 0$$

$$\mu_M(x) = \frac{\text{Sup } f + \text{Inf } f - f}{\text{Inf } f} \quad \text{if Sup } f \leq 0$$

and

$$\mu_M(x) = \frac{f - \text{Sup } f}{\text{Sup } f - \text{Inf } f} \quad \text{if Inf } f \leq 0 \text{ and Sup } f \geq 0.$$

If f is a fuzzy function, that is, if for each $x \in X$, f(x) is a fuzzy set with membership function $\mu_f(x,y)$, then the maximizing set for f(x) is defined by the above equations with f(x) replaced by $\text{Sup}_y \mu_f(x,y)$.

Although the above definitions are precise in character, it should be understood that, in dealing with fuzzy variables, maximization and other operations performed on functions of such variables are highly approximate in nature.

**It should be understood that expressing the total reward as a function of u_t alone is intended merely to single out the dependence of R_t on u_t. In general, R_t will depend, in addition, on the strategy used by H as well as on x_t, s_t, y_t, e_t and possibly other variables.

then uses some random or arbitrary rule to select one among them.

To gain better insight into the operation of the decision principle, it is advantageous to decompose the estimated reward function into two components, one representing an immediate reward or gratification and the other - estimated future reward (or penalty, if the reward is negative). More specifically, we assume that R_t is a function of two arguments: immediate reward function $IR_t(u_t)$ and estimated future reward function $FR_t(u_t)$. Thus, in symbols

$$(7) \qquad R_t(u_t) = G_t(IR_t(u_t),FR_t(u_t))$$

where G_t represents a function* of IR_t and FR_t, playing a role analogous to that of an objective function in control theory. Note that implicit in FR_t is a goal or subgoals in terms of which the consequences of choosing u_t may be estimated.

We are now in a position to make the decription of the behavior of H more explicit by adding to (3), (4), (5) and (6) the equation

$$(8) \qquad u_t = \text{maximizing set for } G_t(IR_t,FR_t)$$

In words, this equation means that H chooses that action u_t which maximizes a specified combination of the immediate reward IR_t and the estimated future reward FR_t, with IR_t and FR_t understood to be known functions of the actions. It should be remarked that the description of the behavior of H by (3),(4),(5),(6) and (8) is consistent with the point of view taken in Skinner's work.

If the variables appearing in equations (3),(4),(5),(6) and (8) were

*As in the case of R_t, it is tacitly understood that G_t may depend on x_t,s_t,y_t, t and possibly other variables.

assumed to be numerically-valued, the task of characterizing the functions $h_1, h_2, g_1, g_2, IR_t, FR_t$ and G would be impossibly complex. The crux of our idea is to regard the variables in question as fuzzy variables ranging over labels of appropriate fuzzy sets.* Equations (3)-(8), then, would represent approximate (that is, fuzzy) relations between fuzzy variables. These relations could be characterized by (a) tables in which the entries are labels of fuzzy sets, or (b) algorithmically, that is, by a set of fuzzy rules (like a computer program with fuzzy instructions) for generating a fuzzy set from other fuzzy sets. In this way, the description of the relations between the variables characterizing human behavior could be greatly simplified - at the cost, of course, of a commensurate loss in precision. In this perspective, the approach sketched above may be viewed as a systematization of the conventional verbal characterizations of human behavior.

When human behavior is described by equations of the form (3),(4), (5),(6) and (8), a modification in human behavior may be viewed as a change in the functions h_1, h_2, G_t, IR_t and FR_t. Of these, the changes in G, IR_t and FR_t play a particularly important role because they influence in a direct way the choice of actions taken by H. Thus, in terms of these functions, Skinner's operant conditioning may be regarded as a form of modification of behavior resulting largely from a manipulation of IR_t through its dependence on the environment.

To clarify the role played by FR_t in relation to IR_t, it will be

*It is understood that the fuzzy sets in question would, in general, be defined in an approximate fashion by exemplification (i.e., ostensively). For example, the fuzzy set <u>very likely</u> would be defined by a collection of examples of probability values together with their grades of membership, e.g., {(1,1.0), (0.9,0.9), (0.8,0.7), (0.7,0.4), (0.6.0.1)}, in which the first element is a probability value and the second element is its grade of membership in the fuzzy set <u>very likely</u>.

convenient to make a very rough approximation to G_t by a numerically-valued convex linear combination

(9) $$R_t = \alpha\, IR_t + (1-\alpha) FR_t$$

in which α is a weighting coefficient, $0 \leq \alpha \leq 1$. Thus, (9) signifies that the reward at time t is a weighted linear combination of the immediate reward and the estimated future reward at time t, with the latter multiplied by the factor $\rho = (1-\alpha)/\alpha$ in relation to the former.

Though not a constant, the anticipation coefficient ρ constitutes an important personality parameter of an individual. In this connection, it should be noted that, in a given individual, ρ will be small when the uncertainty in the estimate FR_t is large. To put it another way, the influence of the immediate reward tends to be predominant when there is considerable uncertainty about the future consequences of an action.

As an individual matures and learns from his own experience as well as that of others, his knowledge of the IR_t and FR_t functions improves and his anticipation coefficient tends to increase, that is, he tends to become more far-sighted. Nevertheless, it is probably true that, judged over a long period of time, the ρ of most individuals is not as large as it should be for their own good as well as the good of others. The acceptance of this premise naturally raises the troublesome question: To what extent should society attempt to coerce its members to increase their anticipation coefficient if they are unwilling to do so on their own volition? Obviously, it is this question that is at the heart of problems relating to smoking, drinking, drug-taking, etc.

It is important to observe that the effect of increasing ρ (for negative FR_t) can also be achieved, for fixed ρ, by decreasing IR_t.

In other words, if an individual tends not to give sufficient weight to long-term harmful consequences of an action which gives him immediate pleasure, then one way of inducing him to modify his behavior is to make IR_t sufficiently negative by adding to it an immediate penalty. For example, one possible way of controlling affinity for excessive drinking might be to implant an electronic monitor in a person who is in need of external reinforcement of his will power. Such a monitor could be programmed to produce an acute sensation of pain or some other form of discomfort when the level of alcohol in blood reaches a predetermined threshold. In this way, the immediate pleasure derived from having one or more drinks would be offset by the nearly simultaneous feeling of pain, with the net immediate reward becoming negative when the amount of alcohol consumed exceeds a set limit.

Behavior-modifying monitors of this type are within the reach of modern electronic technology. Clearly, the potential for abuse of such devices is rather high, for through remote signalling they could be used by a totalitarian government as a highly effective means of punishment and control.

The _temporal_ decomposition of the reward function into two components, one representing the immediate reward and the other - estimated future reward, serves to exhibit an important facet of the decision-making process, namely, the way in which an individual, H, balances short-term gains against long-term losses. In a similar way, we can perform what might be referred to as a _relational_ decomposition of the reward function into components which represent the rewards to other members of a group of individuals who interact with H. Specifically, suppose that we have a group of N individuals $H^1,\ldots,H^N$, with the reward function and action associated with H^i denoted by $R_t^i(u_t^i)$ and u_t^i , respectively.

As a very rough approximation, we assume that $R_t^i(u_t^i)$ admits of the following decomposition*

$$R_t^i(u_t^i) = w_{i1} R_t^{i1}(u_t^i) + w_{i2} R_t^{i2}(u_t^i) + \dots + w_{iN} R_t^{iN}(u_t^i) \tag{10}$$

where

$R_t^{ij}(u_t^i)$ = reward accruing to H^j at time t as a result of action u_t^i taken by H^i

w_{ij} = weight attached by H^i to the reward accruing to H^j as a result of action u_t^i , with $w_{i1} + w_{i2} + \dots + w_{iN} = 1$, $0 \leq w_{ij} \leq 1$

and

$R_t^{ii}(u_t^i)$ = self-reward
= reward accruing to H^i at time t as a result of the action u_t^i taken by H^i.

The basic assumption underlying (10) is that the behavior of H^i is governed not only by the self-reward function R_t^{ii}, but also by a weighted combination of the rewards accruing to other members of the group as a result of the action taken by H^i . More precisely, this implies that when H^i is faced with a decision, he chooses that u_t^i which maximizes R_t^i , as expressed by (10), rather than that u_t^i which maximizes R_t^{ii} .

As in the case of the anticipation coefficient ρ , the relational coefficients $w_1,\dots,w_N$ constitute important parameters of an individual's behavior and personality. In what way does an individual weigh the reward to himself in relation to the rewards to his family, close relatives,

*As in the case of (7), implicit in $R_t^i(u_t^i)$ is the possibility that R_t^i may depend on other variables and actions in additon to u_t^i .

friends, enemies, co-workers, members of the same religion, residents of his community, fellow countrymen etc? Clearly, the answer to this question would be very different for a typical member of a primitive society than for a person of high level of culture and enlightenment. Indeed, the evolution of a society is directly related to the changes in the relational coefficients of its members, with an individual learning from his own experience as well as that of others, that it is in his long-term self-interest to assign greater weight to the interests of not only those who are close to him, but also those who are remote.

In essence, then, once the reward functions IR_t, FR_t and R_t^{ij} have been identified, the behavior modification would involve, in the main, changes in the anticipation coefficient ρ and the relational coefficients w. In the past, changes in ρ and w were induced primarily by experience, education, religious training, political indoctrination and other environmental influences. As implied by Skinner, the time is coming, if if has not come already, when the society will have much more effective means at its disposal for manipulating the ρ and w of its members, perhaps electronically or through systematic psychological conditioning on a mass scale.

To give a simple example of electronic manipulation in a small group, consider a group comprising just two members: H^1 = husband and H^2 = wife. Suppose that each has a device with a push-button such that when the button is pressed, the other party experiences acute pain or discomfort induced by a probe implanted in or attached to the body. Thus, if H^1, say, takes an action which makes H^2 unhappy, then H^2 can retaliate by pressing her button, and vice-versa. To limit the extent of retaliation, both H^1 and

H^2 have a quota which varies from day to day in a random fashion and is not made known to H^1 or H^2. This rule is intended to induce H^1 and H^2 to use their push-buttons rather sparingly.

The point of this example is that the availability of means of retaliation is likely to have the effect of increasing the values of relational coefficients w_{12} and w_{21} in the reward equations

$$(9) \quad R_t^1(u_t^1) = w_{11} R_t^{11}(u_t^1) + w_{12} R_t^{12}(u_t^1)$$

$$(10) \quad R_t^2(u_t^2) = w_{21} R_t^{21}(u_t^2) + w_{22} R_t^{22}(u_t^2)$$

which govern the behavior of H^1 and H^2. However, excessive retaliatory capability or its misuse may, of course, result in a rupture of the relationship between H^1 and H^2.

The use of electronic rather than some other means of retaliation in the above example is intended merely to make retaliation more convenient to apply and hence more effective as a modifier of behavior. The basic point, however, is that whether in small groups or large, the threat of retaliation plays an essential role in tending to increase the values of those relational coefficients which would be small in the absence of retaliatory capability. This is particularly true of the modern technologically-based society, in which the degree of communication and interdependence between distant individuals and groups is far greater than it was in the past.

In the case of inanimate systems, it is an experimentally observed fact that as the degree of interaction (feedback) between the constituents of a system increases, the system eventually becomes unstable. The same phenomenon may well be at the root of the many crises confronting modern society, particularly in race relations, pollution, mass transit, health care, power distribution, monetary systems, employment and education. These crises seem to grow in number and intensity as the technology - in the form of TV, radio, telephone, communication satellites, computers, data banks, jumbo jets and the automobile - rapidly increases the degree of interaction between individuals, groups, organizations, societies and countries. The "culprit" may well be the very basic and universal human desire for freedom, which makes it distasteful for most of us to accept the degree of control and discipline which is needed to maintain societal and interpersonal equilibrium in the face of rapid growth in the degree of interdependence brought about by technological progress. Thus, we are witnessing what may be called the crisis of undercoordination - a crisis which, in the main, is a manifestation of insufficient planning and control in relation to the extent of interaction between the constituents of our society.

Thus, we may be faced with the necessity to curtail our freedoms - perhaps rather extensively - in order to achieve survival in a technologically-based, highly interdependent world of tomorrow. Perhaps this is the crux of Skinner's thesis in "Beyond Freedom and Dignity."

In conclusion, it is quite possible that deliberate, systematic, mass-scale behavior modification employing Skinnerian techniques of operant conditioning, electronic monitors, computers, brain-function altering devices and other paraphernalia of modern technology may become a reality in the not-distant future. I, for one, do not look forward to that day.

REFERENCES

1. B. F. Skinner, "Beyond Freedom and Dignity," A. Knopf, New York, 1971.
2. L. A. Zadeh, "Toward A Theory of Fuzzy Systems," in _Aspects of Network and System Theory_, N. DeClaris and R. Kalman (eds.), Holt, Rinehart and Winston, New York, 1971.
3. R. E. Bellman and L. A. Zadeh, "Decision-Making in a Fuzzy Environment," _Management Science_, Vol. 17, pp. B-141 - B-164, December, 1970.